THE INFLUENCE OF HETEROGENEOUS POROSITY ON SILICON NITRIDE/STEEL WEAR IN LUBRICATED ROLLING CONTACT

J. Kang and M. Hadfield

Bournemouth University, School of Design, Engineering & Computing,
Tribology Design Research Unit, Studland House, 12 Christchurch Road,
Bournemouth, Dorset, BH1 3NA, United Kingdom.
Tel. +44 (0)1202 503750, Fax: +44 (0)1202 503751

Abstract: Heterogeneous porosity is detected on the surface and subsurface of hot isostatically pressed (HIPed) silicon nitride spherical rolling elements. The extent of the localised porosity accounts for an area of six percent of the rolling element surface and four percent of the material volume. An experimental investigation using a rotary tribometer is described to compare the lubricated rolling wear mechanisms and performance of HIPed silicon nitride with heterogeneous porosity defect in contact with steel. A brief review of previous investigations is presented. Localised porosity detection using white and violet light microscopy with post-test evaluation is described. Discussions, micro-hardness measurements and scanning electron microscopy illustrations are presented. Critical localised porosity size is evaluated from experimental results.

Key Words: Si₃N₄ (D), Porosity (B), Defects (B), Fatigue (C), Hybrid bearings
1. INTRODUCTION

1.1 A brief review of silicon nitride wear in lubricated rolling contact

The use of silicon nitride as a material to compete with bearing steels has been considered over the past decades. An important criterion in selecting a hot isostatically pressed silicon nitride ceramic for this application is surface fatigue failure mode. If fatigue failure does occur, typical bearing steel will fail by spalling and this is a non-catastrophic type of failure. Silicon nitride also exhibits a spalling failure mode in addition to zirconia whilst other ceramics such as silicon carbide and alumina fail by fracture. One advantage of using hot isostatically pressed silicon nitride compared to zirconia is the higher rolling contact fatigue life \(^1\). The spalling failure mode is a form of macroscopic wear caused by a fatigue mechanism usually initiated at the position of maximum shear stress within the material subsurface. It should be noted that this type of wear does not depend on direct solid-solid contact although the position of maximum subsurface shear stress will change if friction is involved. Other types of wear have been classified as microscopic (asperity) fatigue, delamination, adhesion, abrasion, corrosion, fretting, erosion and cavitation etc. These types of wear are not mutually exclusive and are classified for convenience of addressing the complex issues involved. Although sliding wear of silicon nitride has been the subject of extensive research programs, the lubricated rolling wear mechanisms have not received such attention. In this study, a brief review of lubricated rolling wear mechanisms will be presented for silicon nitride with some experimental findings and analysis.

The wear behaviour of silicon nitride in lubricated rolling contact was studied experimentally by Scott et al \(^2\) in the early 1970's using a modified four-ball machine. Both reaction-bonded and hot-pressed silicon nitride ball specimens were found to have very poor wear resistance compared with traditional bearing materials ie. En31 steel, M50 steel and tungsten carbide. At that time the micro-hardness varied within the reaction-bonded silicon nitride specimens from 1200 to 800 kg/mm\(^2\) at the surface and centre respectively whilst the hot-pressed specimens were constant at 1300 kg/mm\(^2\). The relationship between microstructure and porosity with the wear performance was discussed. Here a
particular fine microstructure and reduced porosity were identified as desirable. Later Scott and Blackwell ³ studied hot-pressed silicon nitride under conditions of heavily loaded, lubricated, unlubricated and elevated temperatures. Examples of silicon nitride ball samples shattering and multi-phase brittle fractures were found. They concluded that the material was not an effective solution for heavily loaded rolling bearings. By the early 1980's Morrison et. al.⁴ had tested Hot Isostatically Pressed (HIPed) silicon nitride hybrid bearings using 12.7 mm diameter rolling elements, M50 steel rings and a brass cage. Spalling was identified as the only mode of failure and in no cases did the silicon nitride balls fracture. Statistical analysis showed that the bearing life varied according to a power law model. In this case, the silicon nitride was manufactured by the hot-isostatically pressed route. Thus, the evolution of silicon nitride manufacturing processes from reaction-bounded to hot pressed, then to hot isostatically pressed methods illustrated their effects on the rolling wear mechanisms.

A fundamental study into wear of lubricated rolling contacts was reported by Spikes et al in the mid 1980's ⁵. Although this study was performed on metallics its findings are of generic importance to the study of rolling contact wear mechanisms. They reported that wear in lubricated rolling contacts could be controlled by good Elastohydrodynamic (EHD) practice or by traditional policies on hardness control. In this case the wear mechanisms were reported as i) plastic indentation on an asperity scale, ii) fatigue cracking and iii) particle formation by possibly extrusion and fracture near the mouth of the fatigue cracks. It was also reported that this rolling wear behaviour was demonstrably entirely different from that of sliding wear. In an experimental study, Braza et al ⁶ studied the rolling wear behaviour of hot pressed, pressureless sintered and reaction bonded silicon nitride in contact with cast iron. The wear rate for the reaction-bonded material was found to be high in comparison with the other manufacturing methods. Here wear mechanisms were found to be grain pullout for reaction bonded and sub-micron chipping in the case of hot pressed and pressureless sintered silicon nitride. Sliding and rolling wear of ceramics was reviewed by Fischer (1990) ⁷. Two main issues, the relationship between material properties and the mechanical forms of wear, and the importance of tribochemical effects were
discussed. It was emphasised that the chemical nature of the ceramic must be considered in the formulation of lubricants.

More recently, the effects of material composition and physical properties on silicon nitride wear and fatigue performance were studied by Allen. In this paper, it was explained that wear models based solely on hardness and fracture toughness do not correlate with experimental wear and fatigue data. Brittle fracture and tribochemical reactions were found to be prominent mechanism of material removal.

It is clear that extensive experimental studies of silicon nitride wear behaviour in rolling contact are needed to form the theoretical models which can reveal the rolling contact wear mechanism.

1.2 Present study

In this study, lubricated rolling contact four-ball tests are performed on 12.7mm diameter HIPed silicon nitride ball samples which contain localised heterogeneous porosity defects at ball surface and near surface found from balls as procured from a manufacturer. The porosity and wear performance of silicon nitride was investigated in the early 1970s. Since then the manufacturing technology of silicon nitride has advanced, both in material composition and processes. At present, there are two approaches to increase the density and reduce the porosity. One is a compositional approach, which involves additive selection, impurity effects, and phase equilibrium studies. The other is micro-structural approach, which relies on starting material characteristics and optimum selection of process parameters. Porosity of silicon nitride can be largely eliminated through the above mentioned two approaches. Presently hot isostatically pressed silicon nitride is widely used in advanced heat engines, bearings, metal processing, etc. The hot isostatically pressing produces a high-density grade silicon nitride that can endure a high contact pressure for rolling bearing application. Although HIPed silicon nitride commonly approaches a theoretical full density, there is usually some, e.g., 0.01% to a few tenths of one percent, residual porosity (often heterogeneous). This kind of heterogeneous
porosity within the product occurs often in the surface or near surface region. This defect is difficult to
detect during high volume production; hence the study on its influence on rolling contact wear
performance is needed.

Before test, the specimen ball surfaces are examined by microscope under white and violet light
using a dye penetration technique. Micro-hardness at different areas of the specimen ball surfaces are
measured and compared with other steel and ceramic balls (Figure 1). Specimen ball subsurface
investigations are conducted after sectioning and polishing. Scanning Electron Microscopy (SEM) is
employed to observe the microstructure of the localised heterogeneous porosity, normal area, and the
interface region. Accelerated rolling contact fatigue tests are performed using a rotary tribometer at
different loads and lubricants. Contact stresses and minimum film thickness values during test are
calculated (Table 1). The ball surfaces after rolling experiments are examined by a light microscope
using a dye penetration technique and by Talysurf series surface profiler to measure wearscar areas. The
wear scar areas are also observed by SEM.

The purpose of this study is to investigate to what extent these localised heterogeneous
porosity defects affect the wear and fatigue performance of silicon nitride, and to examine the effect of
different lubricant and lubrication condition on these heterogeneous porous microstructure of silicon
nitride under lubricated rolling contact.

2. EXPERIMENTAL PROCEDURE

2.1 Test Machine

Accelerated rolling contact fatigue tests were performed using a PLINT TE92/HS 10,000
rev/min Microprocessor Controlled Rotary Tribometer (Figure 2). This machine is suitable for a variety
of tribology test procedures. In this study, it is configured as a high-speed modified four-ball machine
according to the Institute of Petroleum (UK) IP 300 rolling test procedure. Figure 3 is the schematic of
the modified four-ball rolling configuration. The upper ball is held in the collet and rotates at spindle
speed. The four-ball rolling kinematics was described by Tourret and Wright\textsuperscript{11}. In this study, the upper ball is the specimen ball and three lower balls are standard steel testing balls (specification: 0.5” ball Reference RB12.7/310995A, material: AISI 52100 bearing steel). Through a computer interface, the test speed, load, temperature, etc. are controlled by editing values, ramps and steps in the menu of a TSF computer file. The load is applied by a pneumatic actuator from a supply pressure of 8 bar. The contact region between the upper ball and lower balls is immersed with lubricating oil during each test. The tests are suspended at a set number of drive-shaft revolutions measured by a counter. A vibration sensor automatically stops tests at a pre-determined potentiometer adjustment sensitivity (801) and light emitting diode duration. If upper ball or lower ball failed during testing, machine will vibrate. When the machine vibrates such that the diode remains in the on mode for more than one second, the drive motor will cut out and the timer will stop. This machine can test materials and lubricants at shaft speeds between 100 and 10,000 rev/min. All of the present tests were conducted at a shaft speed of 5000 rev/min.

2.2 Testing Silicon Nitride Rolling Elements

Each testing specimen is a finished 12.7mm diameter HIPed silicon nitride ball, which contains localised heterogeneous porosity defects. The localised heterogeneous porosity defect is difficult to detect during high volume inspection, but it is possible to detect after dye-penetration using a microscope and violet light source. Figure 4 is an overview of a typical specimen ball (B2). The diameter of localised heterogeneous porosity defect is 6 mm; the surface area is approximately 30 mm\textsuperscript{2} that accounts for 6% of the total ball surface. Rolling contact path during the test is through this area and normal area.

Figure 5 illustrates the features of heterogeneous porosity within a typical silicon nitride ball. Figure 5(a) is the overview of the section; the lower left area of the section is the localised heterogeneous porosity defect area. The diameter of the defect is 8.4 mm, the surface area approximately 63.5 mm\textsuperscript{2} which accounts for 12.5 percent of the ball surface. The volume of the defect
is approximately 46.8 mm³, which accounts for 4 percent of the total ball volume (Figure 5(b)). Figure 5(c) and 5(d) show the interface region between localised heterogeneous porosity defect area (lower left) and normal area (upper right), which forms a sharp contrast to each other.

Figure 6 illustrates the SEM comparison of the microstructures within localised heterogeneous porosity defect and normal areas at different magnifications. There is also some porosity in the normal area. Figure 6(a) and 6(d) are at the magnification of 160×. These two sections were compared with ISO 4505-1978(E) Hardmetals—Metallographic Determination of Porosity and Uncombined Carbon

13. Figure 6(a) is close to B08 which is type B apparent porosity at magnification 100×, stands for 4000 pores/ cm², 0.6% porosity in volume. Figure 6(d) is close to A08 at magnification 100× and 200× (0.6% porosity in volume). The only difference between the localised heterogeneous porosity defect area and normal area is the shape and size of the pores. In ISO 4505-1978(E), the pore size is defined as the maximum dimension of the pore in the section. A series (including A08) is used to assess pores with sizes up to 10 µm and B series (including B08) is used to assess pores with sizes ranged from 10 µm to 25 µm. In order to distinguish the shape of a pore, we define the minimum dimension in all directions of a pore as the diameter of pore which is different from the pore size. In the localised heterogeneous porosity defect area, the diameters of pores are most from 4 to 20 µm. In normal area, the diameters of pores are usually less than 2 µm. To characterise surface porosity defects, a dye-penetration method and microscope with ultra-violet light source was utilised to enhance the detection of surface porosity. Figure 7(a) shows a white-light microscopic image of a silicon nitride ball surface; left side is localised heterogeneous porosity defect area. Figure 7(b) shows the same area under ultra-violet light, after dye-penetration treatment.

Surface hardness of different areas (defect area and normal area) on the specimens was measured using a Leitz Miniload 2 micro hardness tester. This tester can measure Vickers, Knoop and scratch hardness with various loads. In this study Vickers hardness was measured using a load of 0.5kg (4.903 N). Measured results show that the hardness in the localised heterogeneous porosity defect areas is quiet low. Figure 1 shows a comparison between the hardness of localised heterogeneous porosity
defect area and normal area of the specimen balls, as well as other brand ceramic balls, and steel balls. Three points on the localised heterogeneous porosity defect area were measured, the hardness of these points even lower than steel balls, ranging from one third to one half of the hardness value of the normal area.

2.3 Lubricant and Lubrication Regime

The effect of the lubricant and lubrication regime on the rolling contact fatigue performance of these silicon nitride is a concerns of this study. Two different lubricants were chosen, one low viscosity lubricant (LV, synthetic oil), the other high viscosity lubricant (HV, a mineral oil). The film thickness between the contact surfaces and the lubrication regime are calculated using the classical Elastohydrodynamic (EHD) equation (1) developed by Hamrock and Dowson\textsuperscript{14} and described for this test arrangement by Hadfield\textsuperscript{15}.

\[ H_{\text{min}} = 3.63 U^{0.68} G^{-0.49} W^{0.68} (1 - e^{-0.68k}) \]  

Here \( U, G, W \) and \( K \) are dimensionless parameters for speed, materials, load and ellipticity respectively. The minimum film thickness (\( h_{\text{min}} \)) is found from the product of the dimensionless minimum film thickness parameter (\( H_{\text{min}} \)) and effective radius (\( h_{\text{min}}=H_{\text{min}}R_x \)). The lubrication regime (lambda) is found from the ratio of minimum film thickness to composite surface roughness (\( \lambda = h_{\text{min}}/\sigma \)), where

\[ \sigma = \sqrt{(R_{q1}^2 + R_{q2}^2)} \]

\( R_{q1} \) and \( R_{q2} \) are the r.m.s. roughness values for the two contacting surfaces. According to Elastohydrodynamic Lubrication theory, if \( \lambda > 3 \), a full fluid film will separate the two surfaces, if \( 1<\lambda<3 \), some asperities contact will occur.

Table 1 summarises the calculated minimum film thickness and lambda ratio for the upper-ball/lower-ball contacting surfaces. The most important finding from these calculations is that the lambda ratios under different load for the LV lubricant are between 2.56 to 2.36, this means that a “mixed” EHD lubrication regime exists during testing at these conditions and hence some asperity
contact will take place. In the case of the HV lubricant ratios under different load, lambda values from 3.24 to 2.99 are calculated and hence full film separation exists and almost no asperity contact is envisaged in theory.

2.4 Contact Stresses

The contact stresses at the upper-ball/lower-ball interface are calculated using Hertz theory and also presented in table 1. Hertz elastic contact stress theory is appropriate for lubrication conditions outlined in section 2.3. The rolling friction coefficient was measured at less than 0.01 and therefore can be neglected for contact stress calculation purposes. The contact load \( P \) (N) between upper ball and each of the three lower balls is calculated using the pneumatic actuator applied load \( P_a \) (kN) and contact angle of 35.3°, shown in equation (3)

\[
P = \frac{1000 \, P_a}{3 \cos \theta}
\]

Hertz contact formulae are used to calculate stress magnitudes, stress positions and contact circle sizes. For a non-conforming spherical contact the maximum compressive stress \( (P_0) \) is given in equation (4) where \( E_1 \) is an effective elasticity modulus for ceramic/steel contact and \( R_1 \) is an effective geometric radius. Equations for contact radius, maximum shear/tensile stresses and stress positions may be obtained in reference books eg.

\[
P_0 = \left[ \frac{6PE_1^2}{\pi^3R_1^5} \right]^{\frac{1}{3}}
\]

The maximum compressive stress of 5.55 GPa is high compared with rolling element bearing operating conditions of around 3.0 GPa but are acceptable for accelerated fatigue testing. Table 1 is the summary of calculated lubrication and contact stresses conditions of each test.
3. POST TEST ANALYSIS AND DISCUSSION

Two test series were performed, using a low viscosity lubricant (LV, synthetic oil) and a high viscosity lubricant (HV, mineral oil), under three different contact loads (160 N, 320 N, 480N) at a constant speed 5000 rev/min. After each test, the ball surfaces were examined using a Talysurf series surface profiler and the wear scar areas calculated using a graphical integration method. Figure 8 summarises the test results in graphical form.

All wear scars occurred at the localised heterogeneous porosity area and little or no wear found at normal area, regardless different contact stresses and different lubricants (figure 9). Because there is some porosity within the normal area (described in section 2.2), it follows that there is a critical pore diameter above which severe wear will occur. In this HIPed silicon nitride under rolling contact, the critical pore diameter is approximately 4 µm. There is an allowable upper limit of residual pores under which it does not affect the wear and fatigue performance very much under lubricated rolling contact. In this case, the allowable diameter of pore is 2 µm. This is in conformation with a previous disk-on-rod rolling contact study which showed that HIPed silicon nitride rod containing micro-porosity performed much better than bearing steel rod[20].

The contact stresses are the dominant factors in rolling contact wear in the localised heterogeneous porosity defect area. For both test series 'A' and 'B', when contact stresses were “low” i.e. maximum compressive stress 3.85GPa and maximum shear stress 1.19GPa, the tests were suspended after the test run reached the pre-set number of revolutions, and the wear scar areas were small (23,117µm² and 20,883µm² respectively). When contact stresses were “high” i.e. maximum compressive stress 5.55GPa and maximum shear stress 1.72GPa, the test times were 13.8 s and 11.9 s, and the wear scar areas were larger (91,182µm² and 126,083µm² respectively). This is because within the localised porosity defect area, the mechanical properties of the material are in the transition from “brittle” to “ductile”. The hardness variation (Figure 1) corresponds to a progressive decrease in elastic modulus, thermal conductivity and flexural strength with increasing amounts of porosity in microstructure[21].
At the “high” contact stresses, subsurface compaction and deformation occurs. Under cyclic rolling contact loading, this will cause irregular surface bulging\textsuperscript{22} and which will produce much higher value of surface roughness than nominated. This will in turn affect the lambda ratio; the actual lambda ratio would be much less than calculated which means there are some asperity contact. At these asperity contact regions, pores cause high stress concentration, microcracking will occur along pore grain boundaries\textsuperscript{23}. This cracking is governed mainly by the shear stress beneath the contact region, featuring distributed shear and compression driven subsurface damage and deformation. The rolling wear mechanism is illustrated in Figure 10 which shows the formation and removal process of material debris in the pore, and notice that in the upper right of the same pore, another debris is almost being formed. Under heavy cyclic loading, this subsurface compaction and deformation, surface bulging, microcracking and material removal process will repeat. Figure 11 is the overview of the wear track in porous area of test 'B1' ball, some material debris within pores is observed.

Figure 12 shows light microscopic images of a silicon nitride ball surface after test, (a) under white-light and (b) the same area under violet light, after dye-penetration treatment. This post-test illustration compares to Figure 6 which is the same area before the test. The wear track is wider in the localised heterogeneous porosity defect area and the irregular emission of light illustrates the variation of the localised porosity. Figure 13 shows a typical wear scar (from Test A2), which has occurred in the localised heterogeneous porosity defect area. No wear scar is observed in the normal area and from figure 13(b) and (c) the wear scar shrink edge is observed. There is no evidence of delamination type fatigue failure, i.e., no cliff edges or fatigue failure dependency of rolling direction as found in previous studies\textsuperscript{24}.

4. CONCLUSIONS

The dye-penetration method and light microscopy with a violet light source can enhance the detection of surface and subsurface localised heterogeneous porosity defects within HIPed silicon nitride. The critical diameter of residual pore is 4µm above which severe wear will occur in lubricated rolling
contact. The allowable diameter of residual pore is 2 µm under which it does not affect the wear and fatigue performance. HIPed silicon nitride with localised heterogeneous porosity defects is extremely susceptible to rolling contact the stress level. In high stress applications, localised heterogeneous porosity defects must be fully eliminated. Compared with low viscosity lubricant, high viscosity lubricant is preferable for silicon nitride with localised heterogeneous porosity defects.

Acknowledgements

The authors of this paper would like to thank the financial support from the SKF Engineering & Research Centre B. V., (The Netherlands) and in particular the technical advice from Dr R. Cundill.
References


<table>
<thead>
<tr>
<th>TEST</th>
<th>CONTACT LOAD (N)</th>
<th>MIN. FILM THICKNESS (m)</th>
<th>LAMBADA RATIO</th>
<th>MAXIMUM COMPRESSIVE STRESS (GPa)</th>
<th>RADIUS OF CONTACT CIRCLE (µm)</th>
<th>MAX. SHEAR STRESS (GPa)</th>
<th>MAX. TENSILE STRESS (GPa)</th>
<th>DEPTH OF MAX. SHEAR STRESS (µm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>480</td>
<td>7.08E-08</td>
<td>2.36</td>
<td>5.55</td>
<td>203</td>
<td>1.72</td>
<td>0.85</td>
<td>97.6</td>
</tr>
<tr>
<td>A2</td>
<td>320</td>
<td>7.29E-08</td>
<td>2.43</td>
<td>4.85</td>
<td>178</td>
<td>1.50</td>
<td>0.74</td>
<td>85.2</td>
</tr>
<tr>
<td>A3</td>
<td>160</td>
<td>7.67E-08</td>
<td>2.56</td>
<td>3.85</td>
<td>141</td>
<td>1.19</td>
<td>0.59</td>
<td>67.7</td>
</tr>
<tr>
<td>B1</td>
<td>480</td>
<td>8.96E-08</td>
<td>2.99</td>
<td>5.55</td>
<td>203</td>
<td>1.72</td>
<td>0.85</td>
<td>97.6</td>
</tr>
<tr>
<td>B2</td>
<td>320</td>
<td>9.23E-08</td>
<td>3.08</td>
<td>4.85</td>
<td>178</td>
<td>1.50</td>
<td>0.74</td>
<td>85.2</td>
</tr>
<tr>
<td>B3</td>
<td>160</td>
<td>9.71E-08</td>
<td>3.24</td>
<td>3.85</td>
<td>141</td>
<td>1.19</td>
<td>0.59</td>
<td>67.7</td>
</tr>
</tbody>
</table>

Table 1: Calculated lubrication conditions and contact stresses

Figure 1: Surface Hardness Comparison Before Test

Vickers Hardness Number

Testing Ceramic Ball (porous defect point1)  Testing Ceramic Ball (porous defect point2)  Testing Ceramic Ball (light grey area)  Testing Ceramic Ball (dark grey area)  Ceramic Ball from Manufacturer A  Ceramic Ball from Manufacturer B  Testing Steel Ball
Figure 2 Plint TE92/HS 10,000 rev/min Rotary Tribometer

Figure 2 Schematic of the modified four-ball machine
1, upper-ball and collet; 2, lower-balls; 3, spindle; 4, pneumatic actuator; 5, drive motor;
6, heated plate; 7, loading piston; 8, belt drive
**Figure 4** Overview of a specimen ball

(a) Overview of section  
(b) Dimension of localised heterogeneous porosity defect (in mm)

(c) Localised heterogeneous porosity defect interface  
(d) Detail of interface region

**Figure 5** Section through a silicon nitride ball (B series)
Figure 6 Comparison of section
(a), (b) and (c) localised heterogeneous porosity defect area; (d), (e) and (f) normal area
a) white light

b) violet light

**Figure 7** Pre-test ball surface ('B' series)

**Figure 8** Summery of test results at 5000 rev/min
Figure 9 Overview of a specimen ball after test

Figure 10 SEM of a pore with debris on wear track from Test 'B1'

Figure 11 SEM of wear track from Test 'B1'
a) white light  

b) violet light

**Figure 12** Post-test ball surface ('B' series)

(a) Overview of localised porous area (← direction of rolling test)

(b) left edge (← direction of rolling test)  

(c) right edge (← direction of rolling test)

**Figure 13** SEM of post-test surface observation